4Н

УДК 681.3

А.А. Никоненко
Киевский национальный университет имени Т. Шевченко, г. Киев, Украина
andrey.nikonenko@gmail.com

Обзор баз знаний онтологического типа

В статье рассматриваются вопросы построения, структурирования, описания, классификации и использования онтологических баз знаний. Приведен обзор современных исследований, посвященных различным аспектам создания и использования онтологии. Пристальное внимание в работе уделено разграничению формальных и лингвистических онтологий. Также предложена достаточно подробная методология построения ресурсов онтологического типа.

Введение

В настоящее время все большей популярностью пользуются методы инженерии знаний, среди которых онтологиям принадлежит значительное место. В связи с ростом популярности изучение онтологий включено в программу ВУЗов, по данной тематике выходит все большее количество статей. К сожалению, на наш взгляд, большая часть русскоязычного материала (в том числе учебные пособия и некоторые статьи) все еще чрезвычайно низкого качества. Как правило, такие материалы содержат части формальных выкладок из переводных источников или разработок отечественных исследований.

Данная статья является попыткой упорядочить и структурировать существующую сегодня информацию о создании и применении онтологий.

1. Понятие онтологии

В исследовании онтологий вопросы возникают с первых шагов. Так, до сих пор не существует единого определения для понятия онтология. Само понятие *онтология* происходит от др.-греч. «онтос» – сущее, «логос» – учение, понятие, т.е. это раздел философии, изучающий бытие.

Википедия определяет *онтологию* (в информатике) как попытку всеобъемлющей и детальной формализации некоторой области знаний с помощью концептуальной схемы [1]. Под концептуальной схемой подразумевается *набор понятий + информация о понятиях* (свойства, отношения, ограничения, аксиомы и утверждения о понятиях, необходимых для описания процессов решения задач в избранной предметной области).

Среди специалистов, занимающихся проблемами компьютерной лингвистики, наиболее устоявшимся (классическим) считается определение онтологии, данное Губертом: «*онтология* – это спецификация концептуализации» [2].

Так же известен ряд расширенных определений Губерта, среди которых можно выделить такие:

– *онтология* – это эксплицитная спецификация концептуализации, где в качестве концептуализации выступает описание множества объектов предметной области и связей между ними;

– *онтология* – это знания, формально представленные на базе концептуализации. Формально онтология состоит из терминов, организованных в таксономию, их определений и атрибутов, а также связанных с ними аксиом и правил вывода;

– *онтология* – формальная спецификация разделяемой концептуализации, которая имеет место в некотором контексте предметной области;
– *онтология* – база знаний, описывающая факты, которые предполагаются всегда истинными в рамках определенного общества на основе общепринятого смысла используемого словаря.

Дополнительно с пониманием термина «онтология» по Губерту можно ознакомиться здесь [3]. Нужно отметить, что определение, данное в [1], полностью передает семантику всех вышеперечисленных определений, поэтому в данной работе будем придерживаться именно его.

2. Виды онтологий

Кроме проблем с точным определением самого понятия «онтология», существует также ряд проблем с описанием модели онтологии формальным языком [4]. Рассмотрим формальное описание онтологии, которое считается классическим:

Онтология определяется как O = <X,R,F>, где
– X – конечное множество понятий предметной области;
– R – конечное множество отношений между понятиями;
– F – конечное множество функций интерпретации [5].

Однако далеко не все существующие сегодня онтологические ресурсы подпадают под приведенное определение. Сегодня эволюция прикладных информационных систем идет в сторону повышения их интеллектуальности. Это оказывает существенное влияние на направление научных и технологических исследований, связанных с использованием компьютеров, а также предоставляет обществу практически значимые результаты. Однако на определенной ступени развития дальнейшее развитие технологий имеющимися средствами становится невозможным. В такие периоды нужен качественный скачок используемых в разработке средств. Одним из таких скачков в области ИИ, направленным на дальнейшую интеллектуализацию систем взаимодействия с пользователем, и явилось появление онтологий.

Поскольку онтологии были ответом науки на нужды своего времени, то и появление их произошло сразу в нескольких областях знания. Соответственно, в каждой из них ресурсы онтологического типа были сформированы по своим, специфичным для области знаний, правилам.

Эту же точку зрения разделяют авторы работы [6], которые считают, что в проектировании онтологий условно можно выделить два направления, до некоторого времени развивавшихся отдельно.

Первое связано с представлением онтологии как формальной системы, основанной на математически точных аксиомах (т.е. это ресурсы онтологического типа, созданные в различных областях математики).

Второе направление развивалось в рамках компьютерной лингвистики и когнитивной науки. Здесь онтология понималась как система абстрактных понятий, существующих только в сознании человека, которая может быть выражена на естественном языке (или средствами какой-то другой системы символов). При этом обычно не делается предположений о точности или непротиворечивости такой системы.

Таким образом, существует два альтернативных подхода к созданию и исследованию онтологий. Первый (формальный) – основан на логике (предикатов первого порядка, дескриптивной, модальной и т.п.). Второй (лингвистический) – основан на изучении естественного языка (в частности, семантики) и построении онтологий на больших текстовых массивах, так называемых корпусах.

3. Классификация онтологий

Существуют различные типы классификации онтологий. С нашей точки зрения, наиболее полезным будет выделить два типа классификации онтологий:

Семантическая:

1) по уровню выразительности [7];
2) по степени формальности [6];
3) по уровню детальности представления [8].

Прагматическая:

4) по степени зависимости от конкретной задачи или прикладной области [9];
5) по языку представления онтологических знаний [9];
6) по предметной области [9];
7) по цели создания [6];
8) по наполнению (содержимому) [6].

Приведем краткую характеристику каждой классификации.

Семантическая классификация

1. По уровню выразительности

Тяжеловесные онтологии. Тяжеловесные онтологии сильно аксиоматизированы, такой уровень аксиоматизации позволяет осуществлять онтологическое связывание явно. Цель аксиоматизации – избежать терминологической и концептуальной неоднозначности из-за неправильной интерпретации. Каждая тяжеловесная онтология может иметь легковесную версию. Многие онтологии предметных областей (domain) – тяжеловесные, т.к. должны поддерживать процессы построения сложных умозаключений. Как и в определении любых размеров, граница между тяжеловесными и легковесными онтологиями размыта.

Легковесные онтологии. Это простые таксономические структуры примитивов или композиции термов с соответствующими определениями. Они слабо аксиоматизированы, т.к. значение термина, используемое внутри комьюнити, обычно более-менее известно всем членам сообщества. Соответственно, связи между терминами, которые рассматриваются как релевантные, могут быть представлены в онтологии ограниченно [7].

2. По степени формальности

Данная классификация схожа с классификацией по уровню выразительности языка описания онтологии (см. предыдущий пункт), однако не эквивалентна ей.

Неформальные. Это онтологии, которые описываются в документе на любом естественном языке (английский, русский, украинский и т.д.). Несмотря на отсутствие формальных правил задания, такие онтологии также могут быть богато наполненными, непротиворечивыми и точными.

Более формализованные. Таксономия может быть двух видов – основанная на термах или на концептах. Такие онтологии, хотя и формализованы, но очень слабо структурированы. Например, структурированные связями «уже чем» или «подкласс» (narrower_than or subclass):

– основанные на термах. В этом случае в иерархии тематического раздела более общие термины находятся выше, по мере спуска по иерархии термины становятся все более специфичными;

– основанные на концептах. Иерархия состоит из классов и их подклассов, в которых отображаются их отличительные и необходимые свойства.

Сильно формализованные. Онтологии для решения инженерных уравнений. Задают формальную семантику термов (таких, как количество и единица измерения) в разрешенных языком точных и непротиворечивых выражениях [8].

3. По уровню детальности представления

Некоторые количественно-измеримые метрики могут быть применены к расчету уровня детальности представления. Например, средняя глубина структуры подклассов/подхарактеристик (subproperty), средняя кучность/кустистость (bushiness), количество аксиом и другие.

Низкий. Онтология может быть построена на основе терминов и нескольких типов связей.

Примеры:

– простая таксономия;

– очень формальная онтология, описанная в терминах общей логики (Common Logic), но содержащая только 3 класса и 2 свойства.

Высокий. Онтология может содержать намного больше деталей, включая правила, по которым термины могут быть связаны между собой.

Примеры:

– очень детализированное описание биологических классов и их отличительных особенностей на естественном языке;

– очень формальная онтология, описанная на одном из языков KIF, CL, Cyc-L, OWL+SWRL, которые содержат тысячи классов, свойств, правил и миллионы экземпляров/индивидуумов [8].

Прагматическая классификация

4. По степени зависимости от конкретной задачи или прикладной области

Верхнего уровня. Такие онтологии описывают наиболее общие концепты (пространство, время, материя, объект, событие, действие и т.д.), которые не зависимы от конкретной проблемы или области. Примером такой общей онтологии является коммерческий проект онтологии CYC.

Ориентированные на предметную область. Во многих дисциплинах сейчас разрабатываются стандартные онтологии, которые могут использоваться экспертами по предметным областям (доменам) для совместного использования и аннотирования информации в своей области. Например, в области медицины созданы большие стандартные, структурированные словари, такие как SNOMED и семантическая сеть Системы унифицированного медицинского языка (the Unified Medical Language System). Также появляются обширные *общецелевые онтологии,* например, онтология UNSPSC, которая предоставляет терминологию товаров и услуг.

Ориентированные на задачу. Это онтология, используемая конкретной прикладной программой и содержащая термины, которые используются при разработке ПО, выполняющего конкретную задачу. Она отражает специфику приложения, но может также содержать некоторые общие термины (например, в графическом редакторе будут и специфические термины – палитра, тип заливки, наложение слоев и т.д., и общие – сохранить и загрузить файл).

Прикладные онтологии описывают концепты, которые зависят как от онтологии задач, так и от онтологии домена. Примером может служить онтология для автомобилей, строительных материалов, вычислительной техники [9].

5. По языку представления онтологических знаний

RDF. Язык разработан в рамках проекта Семантик-веб (Semantic Web). Основное предназначение языка – описание метаданных документов, размещаемых в Интернете. RDF использует базовую модель представления данных «объект – атрибут – значение», и способен сыграть роль универсального языка описания семантики ресурсов и связей между ними.

DAML+OIL – семантический язык разметки Web-ресурсов, который расширяет стандарты RDF і RDF Schema за счет более полных примитивов моделирования. В последнюю версию DAML+OIL включен набор дополнительных конструкций для создания онтологий и разметки информации в легкоинтерпретируемом машиной виде.

OWL (Web Ontology Language) – язык представления онтологий следующего поколения после DAML+OIL. Обладает более богатым набором возможностей, чем XML, RDF, RDF Schema і DAML+OIL. Проект предполагает создание мощного механизма семантического анализа. Планируется, что в нем будут устранены ограничения конструкций DAML+OIL. Онтология OWL – это последовательность аксиом, фактов и ссылок на другие онтологии.

KIF (Knowledge Interchange Format, или формат обмена знаниями) – основан на S-выражениях синтаксиса для логики. KIF – специальный язык, предназначенный для обмена знаниями между разными компьютерными системами. Разрабатывался для описания общего формата представления знаний, независимого от конкретных систем.

CycL (язык описания онтологии Cyc) – это гибридный язык, в котором объединены свойства фреймов и логики предикатов. Синтаксис языка CycL схож с синтаксисом языка Lisp. CycL различает такие сущности, как экземпляры, классы, предикаты и функции. Словарь CycL состоит из термов. Множество термов можно разделить на константы, неатомарные термы и переменные. Термы используются при составлении значащих выражений CycL, из которых формируются суждения. Из суждений состоит база знаний [10].

OCML (Operational Conceptual Modeling Language) язык поддерживает построение нескольких типов конструкций представления знаний. Он позволяет задавать спецификацию и операционализацию функций, связей, классов, экземпляров и правил. Он также включает механизмы для описания онтологий и методов решения задач – основные технологии, разработанные в области представления знаний. Около десятка проектов в KMI (Knowledge Media Institute) в настоящее время используют OCML для разработки моделей в таких областях, как управление знаниями, разработка онтологии, электронная торговля и системы обработки знаний [11].

LOOM и PowerLoom® – языки представления знаний, разработанные исследователями из группы Artificial Intelligence Research Group [12] Университета Южной Калифорнии (University of Southern California's Information Sciences Institute). Цель проекта Loom – разработка и внедрение продвинутых средств для представления знаний и рассуждений в области искусственного интеллекта. Loom и PowerLoom распространяются по открытой лицензии (open source licenses), но являются интеллектуальной собственностью Университета Южной Калифорнии и не являются общедоступными.

Loom – это и язык и среда для построения интеллектуальных приложений. Центром языка является система представления знаний, которая используется для построения дедуктивных выводов на основе декларативных знаний. Декларативные знания состоят из определений, правил, фактов и правил по умолчанию. Дедуктивный движок использует прямые цепочки логического вывода, семантическую унификацию и объектно-ориентированные технологии поддержания достоверности [13].

Ontolingua предоставляет распределенную среду для совместного просмотра, создания, редактирования, изменения и использования онтологий. Сервер поддерживает более 150 активных пользователей, с некоторыми проектами которых можно ознакомиться по адресу http://ontolingua.stanford.edu/doc/ontology-server-projects.html [14]. Ontolingua состоит из KIF парсера, инструментов для анализа онтологии, и набора трансляторов для преобразования исходных данных Ontolingua в форму, приемлемую для внедрения в системы представления знаний [15].

F-Logic – онтологический язык, который базируется на логиках первого порядка, однако классы и свойства в нем представлены как термины, а не как предикаты. Язык создавался для осуществления взаимодействия между онтологиями, построенными на основе предикатов, и онтологиями, построенными на основе F-Logic. Создатели определили интуитивные трансляторы для преобразования знаний из предикатных онтологий в F-Logic онтологии и показали, что такой перевод сохраняет логические связи (preserves entailment) для большого количества онтологических языков, в том числе и для многих OWL DL. Также язык может применяться для метамоделирования расширений Description Logics (v-semantics) [16].

6. По предметной области

Онтология отражает общие знания о предметной области, такие, как иерархия классов понятий и семантические отношения на этих классах. Для каждой предметной области онтологии создаются экспертами своей области, которые проводят формализацию знаний, определений и правил получения новых знаний. Для создания и поддержки онтологий существуют как спецификации, так и инструменты (описаны в [17]).

7. По цели создания

Онтологии приложения. Используются во время выполнения конкретного приложения, осуществляющего онтологическое нанесение ограничений на аксиоматизацию для терминологической службы, т.е. используются в работе блока построения рассуждений. Типичный компромисс между выразительностью и разрешимостью требует ограниченного представления формализмов. В приложениях, построенных на принципах описательных логик (DL), это будет совпадать с TBox. Онтологии приложения могут также описывать конкретные миры (семантические описания, базы знаний, метаданные, семантические метаданные или просто экземпляры). В приложениях на базе описательных логик это будет совпадать с ABox.

Ссылочные онтологии. Используются во время разработки приложений, для взаимного понимания и толкования между агентами, принадлежащими к разным сообществам, для установления консенсуса между сообществами, которые нуждаются во введении нового термина, или просто для пояснения значения термина новому участнику сообщества. Хотя части ссылочной онтологии могут быть также формализованы, как TBox, однако описательные логики обычно не достаточно выразительны, чтоб использовать их в качестве справочных материалов [7].

8. По наполнению (содержимому)

Данная классификация (рис. 1) очень похожа на классификацию по цели создания, однако акцент делается на реальное содержимое онтологии, а не на абстрактную цель, преследуемую авторами [6].

Дополнительно можно ввести и другие классификации.

Например, онтологии могут быть также разделены на одноязычные и многоязычные. Уже существует ряд онтологий, ориентированных на представление знаний на нескольких языках, например, EuroWordNet, MikroKosmos и некоторые другие. Сложность создания таких онтологий обычно заключается в том, что возможно наличие различий в понятийных системах разных языков.

Также все онтологии могут быть разделены на глубинные и поверхностные. Поверхностные онтологии строятся на поверхностной семантике, они определяют понятия через значения слов. Однако здесь возникает проблема, какое количество смыслов выделять для каждого слова. Глубинные же онтологии используют глубинную семантику [18].

Иногда лексические (лингвистические) онтологии включают в классификацию как отдельный тип. Отличительным свойством таких онтологий является «фиксация в одном ресурсе (лексикализованных) понятий (слов) вместе с их языковыми свойст-

вами». Такие онтологии тесно взаимосвязаны с семантикой грамматических элементов (слов, именных групп и др.). Основным источником понятий в онтологиях данного типа являются значения языковых единиц. Их также отличает своеобразный набор отношений, обычно свойственный для языковых элементов: синонимия, гипонимия, меронимия, а также ряд других. К лингвистическим онтологиям авторы относят WordNet, MikroKosmos, Sensus, РуТез и другие. Круг задач, решаемых такими онтологиями, тесно взаимосвязан с обработкой естественного языка [19].

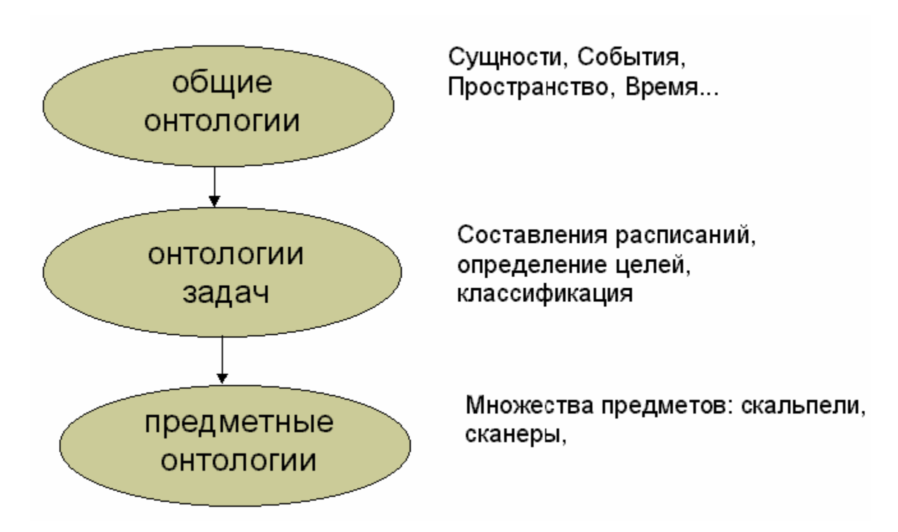


Рисунок 1 – Схема классификации онтологий по наполнению

4. Структура формальных и лингвистических онтологий

Формальная онтология – это система, которая состоит из набора понятий и набора утверждений об этих понятиях, на основе которых можно строить классы, объекты, отношения, функции и теории.

Большинство моделей онтологий содержат такие компоненты:

– *концепты* (понятия, классы);
– *свойства* концептов (атрибуты, роли);
– *отношения* между концептами (зависимости, функции);
– дополнительные *ограничения*, которые определяются аксиомами;
– примеры использования.

В роли концепта может выступать описание задачи, функции, действия, стратегии, процесса соображения и т.д. Основное отличие онтологической системы от обычного словаря – внутреннее единство, логическая взаимосвязь и непротиворечивость используемых понятий.

При построении онтологических систем стараются соблюдать следующие правила:

– *формализации*, т.е. описания объективных элементов действительности в единых, строго определенных образцах (терминах, моделях и др.);
– использования *ограниченного количества базовых терминов* (сущностей), на основе которых конструируются все остальные понятия;
– внутренней *полноты* и логической *непротиворечивости* [9].

Формально эти правила представлены следующими пятью принципами:

1. *Принцип полноты*. Категории верхнего уровня должны исчерпывающим образом представлять Материю; за пределами этих категорий не должно существовать никаких проявлений сущего.

2. *Принцип естественнонаучности и проблемной ориентации*. Все категории и концепты онтологии должны быть выражены понятиями, которые установились в естественных и математических науках при изучении материального мира и являются общепринятыми. При этом часть онтологии должна быть представлена концептами, которые широко используются в междисциплинарных текстах (с нейтральной, общедоступной лексикой), а вторая часть онтологии структурируется под конкретную область знаний (ПрО). Первая часть имеет постоянный статус, а проблемно-ориентированная онтология формируется специалистом и носит переменный характер.

3. *Принцип взаимосвязанности уровней.* Категории онтологии верхнего уровня раскрываются наборами концептов среднего уровня. В свою очередь, концепты нижнего уровня должны служить определителями для терминов словаря ПрО. Связь между средним и нижним уровнями организуется с помощью именованных отношений вида: «быть частью», «принадлежать множеству», «совпадать с», «находиться в семантическом отношении с».

4. *Принцип ассоциативности*. Концепты онтологии нижнего уровня должны служить полем для индексирования терминов ПрО. При этом используются семантические отношения вида: «находиться в ассоциативной связи с».

5. *Принцип отражения антагонизмов*. Концепты, которые отражают свойства или понятия, имеющие свою противоположность или дополнительность по равному основанию, входят в онтологию парами или тройками полярных обозначений [20].

Второй вид онтологий – это иерархические лексические ресурсы типа WordNet. При описании ресурсов этого типа мы будем использовать информацию, представленную в работе [21]. WordNet-ресурсы описывают лексические отношения между значениями слов, представленные в виде отдельных единиц в иерархической сети – *синсетов*. Отношения между лексическими единицами в значительной мере отражают отношения объектов внешнего мира, поэтому такие ресурсы часто рассматриваются как особый вид онтологий – *лексические* или *лингвистические онтологии*.

Главной характеристикой лингвистических онтологий является то, что они привязаны к значениям (are bound to the semantics) языковых выражений (слов, именных групп и т.п.). Лингвистические онтологии охватывают большинство слов языка и одновременно имеют онтологическую структуру, проявляющуюся в отношениях между понятиями. Поэтому лингвистические онтологии могут рассматриваться как особый вид лексической базы данных и особый тип онтологии.

Основное отличие лингвистических и формальных онтологий – это степень формализации. Предполагается, что при разработке ресурсов такого типа строится иерархия лексических значений естественного языка, а для более строгого описания знаний о мире необходимо сопоставить такие ресурсы с какими-либо формальными онтологиями.

Так, содержанием одного из проектов является установление отношений между WordNet и EuroWordNet, с одной стороны, и формальной онтологией SUMO – Standardized Upper Merged Ontology – с другой. Проект состоит в том, чтобы установить соответствие между синсетами WordNet и понятиями онтологии, при котором каждый синсет WordNet либо напрямую сопоставляется с понятием онтологии, либо является гипонимом для некоторого понятия или экземпляром (элементом) понятия онтологии.

Участники другого проекта – OntoWordNet – считают, что недостаточно провести формальную склейку ресурса типа WordNet и формальной онтологии: необходима значительная реструктуризация исходного лексического ресурса. Соответствующая работа проводится в проекте Cyc.

5. Методология построения онтологии

На тему «как построить онтологию» существует множество рекомендаций, руководств, советов, подсказок и указаний. Однако до сих пор не предложено общей методологии построения онтологий. Некоторые шаги в этом направлении были предприняты в следующих работах:

– IDEF5: язык и метод для описания онтологий [22];

– статья Грубера, в которой приведено обсуждение принципов описания онтологий [23];

– статья Скуса, посвященная вопросам согласования, как пути для совместного использования и интеграции онтологий, разработанных разными группами [24];

– статьи Гомез-Переза, посвященные технологиям совместного использования знаний [25].

Указанные выше работы приводят дальнейшие шаги в развитии методологии, но не предлагают базовой (skeletal) методологии. Далее в этом разделе мы попытаемся вывести такую базовую методологию, опираясь на работы [9] и [26]. Такая методология должна включать ограниченное число стадий и быть применимой к построению любой онтологии.

В работе [26] авторы считают, что комплексная базовая методология для разработки онтологий должна включать следующие стадии:

- Определение цели.
- Построение онтологии.
  - Зафиксировать онтологию.
  - Закодировать онтологию.
  - Провести интеграцию существующих онтологий.
- Оценить.
- Задокументировать.

Дополнительно для каждой стадии должны быть описаны технология, методика, принципы и рекомендации, например, типы связей между стадиями (рекомендуемый порядок, чередование, входы/выходы).

Дадим краткую характеристику описанным выше стадиям.

Определение цели

Эта стадия необходима для четкого понимания, почему онтология была создана и для каких целей она будет использована. Она также будет, использоваться для определения круга пользователей онтологии.

Построение онтологии

Фиксирование

Под фиксированием подразумевается:

1. Определение ключевых концептов и связей в нужном домене (scoping).
2. Выработка точных, непротиворечивых текстовых определений для этих концептов и связей.
3. Определение термов, относящихся к таким концептам и связям.
4. Согласование всего вышеперечисленного.

Кодирование

Под кодированием подразумевается точное представление деталей, зафиксированных на предыдущей стадии, на каком-либо формальном языке. Это включает в себя фиксацию определенной мета-онтологии, выбор языка представления, и создание описания онтологии на этом языке.

Интеграция

На этапах фиксирования и кодирования часто возникает вопрос, можно ли как-то воспользоваться уже существующими онтологиями. В общем случае это очень сложная проблема, хотя, например, в проекте Ontolingua был достигнут значительный прогресс в этом направлении. Разработка руководства и инструментария в данной области стало бы одним из наиболее значительных прорывов в области разработки комплексной методологии.

Оценка

Гомез-Перез предложил следующее определение оценки в контексте технологий совместного использования знаний: «для составления технического мнения об онтологии, ассоциированной с ней программной среды и документации, касательно критериев… Критериями могут выступать требования к спецификации, вопросы компетенции, и/или реальный мир» [25]

Документирование

Может быть желательным определить рекомендации к документированию онтологий, возможно, различные, в зависимости от типа и предназначения онтологии. Как показано в работе [24], одним из основных препятствий в эффективном совместном использовании знаний является недостаточная документация по существующим базам знаний и онтологиям. Поэтому все важные допущения должны быть задокументированы, как об основных концептах, определенных в онтологии, так и о примитивах, используемых в определениях.

Как итог, приведем список действий (подробнее в работе [9]), которые необходимо выполнить при построении онтологии:

– обозначение целей и области применения создаваемой онтологии;
– построение онтологии, которое включает:

a) фиксирование знаний о ПрО, т.е. определение основных понятий и их взаимоотношений в выбранной предметной области; создание точных непротиворечивых определений для каждого основного понятия и отношения; определение терминов, которые связаны с этими терминами и отношениями;

b) кодирование, т.е. разделение совокупности основных терминов, используемых в онтологии, на отдельные классы понятий;

c) выбор или разработку специального языка для представления онтологии;

d) непосредственно задание фиксированной концептуализации на выбранном языке представления знаний;

– совместное использование людьми или программными агентами общего понимания структуры информации;
– обеспечение возможности использования знаний ПрО;
– создание явных допущений в ПрО, лежащих в основе реализации;
– отделение знаний ПрО от оперативных знаний – это еще один вариант общего применения онтологий;
– анализ знаний в ПрО.

6. Применение онтологий

Потребность в онтологиях связана с невозможностью адекватной автоматической обработки естественноязыковых текстов существующими средствами. Так, например, создание тезаурусов не решает проблему, т.к. различные группы пользователей и сообщества, занимающиеся обработкой и анализом информации, используют специальную терминологию, которая может применяться другими сообществами в ином контексте. Также в различных сообществах часто встречаются различные обозначения для одних и тех же понятий.

Поэтому для качественной обработки текстов необходимо иметь детальное описание проблемной области с множеством логических связей, которые показывают соотношения между терминами области. Использование онтологий позволяет представить естественноязыковый текст в таком виде, что он становится пригодным для автоматической обработки.

Дополнительно онтологии могут использоваться в качестве посредника между пользователем и информационной системой, что позволяет формализовать используемые термины между всеми пользователями проекта.

Также широкое применение нашли задачи онтологического анализа. В рамках этих задач с помощью онтологического исследования накапливают ценную информацию о функционировании сложных систем. Такой анализ обычно начинается с составления словаря терминов, который используется при обсуждении и исследовании характеристик объектов и процессов, составляющих рассматриваемую систему, а также создания системы точных определений этих терминов. Кроме того, документируются основные логические взаимосвязи между соответствующими терминами и понятиями. Результатом этого анализа является словарь терминов, их точных определений и взаимосвязей между ними. Собранную информацию используют при проведении процесса реорганизации существующих или построения новых систем.

Более подробную информацию по использованию онтологий можно найти в работе [27].

Литература

1. Онтология_(информатика) [Электронный ресурс]. Энциклопедия Википедия. – Режим доступа : http://ru.wikipedia.org/wiki/Онтология_(информатика)
2. Gruber T.R. A translation approach to portable ontologies [Электронный ресурс] / T.R. Gruber // Knowledge Acquisition. 1993. – № 5(2). – С. 199-220. – Режим доступа : http://www-ksl.stanford.edu/kst/what-is-an-ontology.html
3. Gruber T.R. Ontology to appear in the Encyclopedia of Database Systems [Электронный ресурс] / Ling Liu and M. Tamer Özsu (Eds.). – Springer-Verlag, 2008. – Режим доступа : http://tomgruber.org/writing/ontology-definition-2007.htm
4. Онтологии и тезаурусы: модели, инструменты, приложения / [Добров Б.В., Иванов В.В., Лукашевич Н.В., Соловьев В.Д.] [Электронный ресурс] : учебный курс Интернет-университета информационных технологий. – Режим доступа : http://www.intuit.ru/department/expert/ontoth/2/
5. Онтология [Электронный ресурс]. Энциклопедия Википедия. – Режим доступа : http://ru.science.wikia.com/wiki/Онтология
6. Онтологии и тезаурусы : учебное пособие / [Соловьев В.Д., Добров Б.В., Иванов В.В., Лукашевич Н.В.]. – Казань ; Москва, 2006. – 173 с.
7. Core Software Ontology. Core Ontology of Software Components. Core Ontology of Services [Электронный ресурс]. OntoWare Group. – Режим доступа : http://cos.ontoware.org/
8. Gruninger Michael. An Ontology Framework / Michael Gruninger, Leo Obrst // Ontology Summit – NIST, (Gaithersburg, MD April 22 – 23, 2007). – Gaithersburg, 2007.
9. Гладун А.Я. Онтологии в корпоративных системах. Часть I / А.Я. Гладун, Ю.В. Рогушина // Корпоративные системы. – 2006. – № 1. – С. 41-47.
10. Глибовець М.М. Побудова україномовної онтології засобами СУБД / М.М. Глибовець, О.О. Марченко, А.О. Никоненко // Наукові записки. Комп'ютерні науки / Національний університет «Києво-Могилянська академія». – 2008. – Том 86. – С. 46-49.
11. Operational Conceptual Modelling Language [Электронный ресурс]. – Режим доступа : http://technologies.kmi.open.ac.uk/ocml/
12. Intelligent Systems Division [Электронный ресурс]. USC/Information Sciences Institute (ISI). – Режим доступа : http://ai.isi.edu/
13. Loom Project Home Page [Электронный ресурс]. Artificial Intelligence research group University of Southern California's Information Sciences Institute. – Режим доступа : http://www.isi.edu/isd/LOOM/
14. Ontolingua [Электронный ресурс]. Artificial Intelligence Laboratory of the Department of Computer Science at Stanford University. – Режим доступа : http://www.ksl.stanford.edu/software/ontolingua/
15. Ontolingua Reference Manual Ontolingua [Электронный ресурс]. Laboratory for Applied Ontology of the Institute of Cognitive Science and Technology Italian National Research Council. – Режим доступа : http://www.loa-cnr.it/medicine/reference-manual/index.html
16. On the Relationship between Description Logic-based and F-Logic-based Ontologies [Электронный ресурс] / Jos de Bruijn, Stijn Heymans // Fundam. Inform. – 2008. – № 82(3). – С. 213-236. – Режим доступа : http://portal.acm.org/citation.cfm?id=1377800
17. Овдей О.М. Обзор инструментов инженерии онтологий / О.М. Овдей, Г.Ю. Проскудина // Электронные библиотеки. – 2004. – Т. 7, вып. 4. – С. 3-19.

18. Константинова Н.С. Онтологии как системы хранения знаний [Электронный ресурс] / Н.С. Константинова, О.А. Митрофанова. – Режим доступа : http://www.sci-innov.ru/icatalog_new/index.php?action=send_att&entry_id =68352&fname=68352e2-st08_(Митрофанова О.А.).pdf
19. Добров Б.В. Курс из 16 презентаций: «Онтологии и тезаурусы» [Электронный ресурс] / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. – Режим доступа : http://window.edu.ru/window_ catalog/redir?id=41722&file=ot_2006_posobie.pdf
20. Гладун Виктор. Структурирование онтологии ассоциаций для конспектирования естественно-языковых текстов [Электронный ресурс] / Виктор Гладун, Виталий Величко, Леонид Святогор. – Режим доступа : http://www.foibg.com/ibs_isc/ibs-02/IBS-02-p20.pdf
21. Онтологии для обработки текстов на естественном языке. Лексические онтологии [Электронный ресурс]. – Режим доступа : http://www.intuit.ru/department/expert/ontoth/2/3.html
22. KBSI. Knowledge based systems incorporated, Technical Report IDEF5 Method Report [Электронный ресурс] AL/HRGA, Wright-Patterson Air Force Base, Ohio, September 1994. – Режим доступа : http://www.idef.com/pdf/Idef5.pdf
23. Gruber T. Towards principles for the design of Ontologies used for knowledge sharing // International Journal of Human-Computer Studies. – 1995. – № 43(5/6). – С. 907-928.
24. Skuce D. Conventions for reaching agreement on shared Ontologies / D. Skuce // In Proceeding of the 9th Knowledge Acquisition for Knowledge Based Systems Workshop, 1995 [Электронный ресурс]. – Режим доступа : http://portal.acm.org/ft_gateway.cfm?id=1135849&type=pdf&coll=GUIDE&dl=GUIDE&CFID=43461927&CFTOKEN=47811814
25. Gomez-Perez A. Evaluation and assessment of knowledge sharing technology // Towards Very Large Knowledge Bases – Knowledge Building and Knowledge Sharing 1995 / A. Gomez-Perez, N. Juristo, J. Pazos ; in N.J. Mars, editor. – Amsterdam : IOS Press, 1995. – P. 289-296.
26. Uschold Mike. Towards a Methodology for Building Ontologies [Электронный ресурс] / Mike Uschold, Martin King. – Режим доступа : http://www.aiai.ed.ac.uk/project/pub/documents/1995/95-ont-ijcai95-ont-method.ps
27. Гладун А.Я. Онтологии в корпоративных системах. Часть II / А.Я. Гладун, Ю.В. Рогушина // Корпоративные системы. – 2006. – № 1. – С. 48-56.

А.О. Никоненко

Огляд баз знань онтологічного типу

У статті розглядаються питання побудови, структурування, опису, класифікації та використання онтологічних баз знань. Наведено огляд сучасних досліджень, присвячених різним аспектам створення і використання онтології. Особливу увагу в роботі присвячено розмежуванню формальних і лінгвістичних онтологій. Також запропоновано досить детальну методологію побудови ресурсів онтологічного типу.

A.A. Nykonenko

The Ontological Knowledge Bases Review

In the article the issues of construction, structurization, description, classification and use of the ontological knowledge base are considered. The review of the modern researches, dedicated to different aspects of ontology use and formation, is made. Special attention is given to the delimitation of formal and linguistic ontologies. The detailed methodology of ontology resources building is offered as well.

Статья поступила в редакцию 10.06.2009.